DTIC FILE COPY

JAN 1988 2 7

AD-A193 399

(2)

ON PAGE — READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 88-0202	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Asymptotic normality of minimum L_1-norm estimates in linear models		5. TYPE OF REPORT & PERIOD COVERED: ~~Technical~~ Journal - September 1987
		6. PERFORMING ORG. REPORT NUMBER: 87-35
7. AUTHOR(s): Z.D. Bai, X.R. Chen, Y. Wu and L.C. Zhao		8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0292; F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Multivariate Analysis, Fifth Floor Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F/2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research; Air Force Office of Scientific Research NM		12. REPORT DATE: September 1987
		13. NUMBER OF PAGES: 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Same as 11		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE MAR 0 4 1988 S D D

18. SUPPLEMENTARY NOTES

19. *Key words and phrases:* linear model, Minimum L_1-Norm estimate, consistency, asymptotic normality.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider the standard linear model $Y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + e_i$, $i = 1,\ldots,n,\ldots$ where $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots$ are assumed to be known p-vectors, $\underset{\sim}{\beta}$ the unknown p-vector of regression coefficients, and $e_1, e_2, \ldots$ the independent random error sequence each having a median zero. Define the Minimum L_1-Norm estimator

DD FORM 1 JAN 73 1473 — Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (continued)

$\hat{\beta}_n$ as the solution of the minimization problem $\sum_{i=1}^n |Y_i - x_i'\hat{\beta}_n| = \inf\{\sum_{i=1}^n |Y_i - x_i'\beta| : \beta \in R^p\}$. It is proved in this paper that $\hat{\beta}_n$ is asymptotically normal under very weak conditions. In particular, the condition imposed on $\{x_i\}$ is exactly the same which ensuring the asymptotic normality of Least Squares estimate: $\lim_{n\to\infty} \max_{1\le i\le n} x_i'(\sum_{j=1}^n x_j x_j')^{-1} x_i = 0$.

AFOSR-[redacted] 8 8 ...2

Center for Multivariate Analysis

University of Pittsburgh

88 2 23 155

ASYMPTOTIC NORMALITY OF MINIMUM

L_1-NORM ESTIMATES IN LINEAR MODELS*

Z.D. Bai, X.R. Chen, Y. Wu and L.C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 87-35

September 1987

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

ASYMPTOTIC NORMALITY OF MINIMUM L_1-NORM ESTIMATES IN LINEAR MODELS*

Z.D. Bai, X.R. Chen, Y. Wu and L.C. Zhao

ABSTRACT

Consider the standard linear model $Y_i = \underset{\sim}{x}_i'\beta + e_i$, $i = 1,\ldots,n,\ldots$ where $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots$ are assumed to be known p-vectors, $\underset{\sim}{\beta}$ the unknown p-vector of regression coefficients, and $e_1, e_2, \ldots$ the independent random error sequence each having a median zero. Define the Minimum L_1-Norm estimator $\hat{\underset{\sim}{\beta}}_n$ as the solution of the minimization problem $\sum_{i=1}^n |Y_i - \underset{\sim}{x}_i'\hat{\underset{\sim}{\beta}}_n| = \inf\{\sum_{i=1}^n |Y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta}| : \underset{\sim}{\beta} \in R^p|$. It is proved in this paper that $\hat{\underset{\sim}{\beta}}_n$ is asymptotically normal under very weak conditions. In particular, the condition imposed on $\{x_i\}$ is exactly the same which ensuring the asymptotic normality of Least Squares estimate: $\lim_{n\to\infty} \max_{1\le i\le n} \underset{\sim}{x}_i(\sum_{j=1}^n \underset{\sim}{x}_j\underset{\sim}{x}_j')^{-1}\underset{\sim}{x}_i = 0$.

AMS 1980 subject classifications: Primary 62J05; secondary 62F12.

Key words and phrases: linear model, Minimum L_1-Norm estimate, consistency, asymptotic normality.

* This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

1. INTRODUCTION AND SUMMARY

Consider the standard linear regression model

$$Y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta}_0 + e_i, \quad i = 1,\ldots,n,\ldots \tag{1.1}$$

where $\underset{\sim}{x}_1$, $\underset{\sim}{x}_2$, ... are assumed to be known p-vectors, $\underset{\sim}{\beta}_0$ the unknown p-vector of regression coefficients, and e_1, e_2, ... the i.i.d. random errors with a common density function f and median zero, f is continuous at 0 and $f(0) > 0$. The Minimum L_1-Norm (ML_1N) estimate $\hat{\underset{\sim}{\beta}}_n$ of $\underset{\sim}{\beta}_0$ is defined as a solution of the

$$\sum_{i=1}^{n} |Y_i - \underset{\sim}{x}_i'\hat{\underset{\sim}{\beta}}_n| = \inf\{\sum_{i=1}^{n} |Y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta}| : \underset{\sim}{\beta} \in R^p\}. \tag{1.2}$$

Here we assume that the parameter space is the whole p-dimensional Euclidean space R^p. It will be indicated (see Corollary 3 below) that no change in the argument is needed when R^p is replaced by any of its subset containing the true parameter $\underset{\sim}{\beta}_0$ as an inner point.

The ML_1N estimate, whose usefulness is by now universally recognized, dates back to Laplace. But for a long time in history it never attracted much attention. One reason is in the difficulty of its computation, which has now been resolved with the advent of modern computing facilities, and the paper of Charnes *et al* (1955) linking the computation of $\hat{\underset{\sim}{\beta}}_n$ to the solution of a linear programming problem. Another reason is the lack of an adequate asymptotic theory. It is well known that in the problem of estimating the median of a univariate population, the sample median is (under certain conditions) asymptotic normal. Motivated by this simple case, write

$$\underset{\sim}{S}_n = \sum_{i=1}^{n} \underset{\sim}{x}_i\underset{\sim}{x}_i'. \tag{1.3}$$

It is naturally expected that ($\xrightarrow{\mathcal{L}}$ means convergence in distribution)

$$2f(0)\underset{\sim}{S}_n^{1/2}(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}_0) \xrightarrow{\mathcal{L}} N(\underset{\sim}{0}, \underset{\sim}{I}_p), \quad \text{as } n \to \infty \tag{1.4}$$

under reasonable conditions. Here $\underset{\sim}{I}_p$ is the identity matrix of order p.

The first attempt to give a proof of (1.4) was made by Bassett *et al* (1978). They assumed that $\{e_i\}$ satisfies the conditions stated earlier, the solution of (1.2) is unique (a condition difficult to justify), and that

$$\underset{\sim}{S}_n/n \to \underset{\sim}{Q}, \quad \text{a positive definite matrix.} \tag{1.5}$$

Unfortunately their argument contains serious mathematical gaps which do not seem easy to resolve. For one thing, they overlooked the fact that the o(1) at the right-hand side of the equation above (3.10) of their paper should be $o_h(1)$, and it is by no means clear that the convergence (as $T \to \infty$) $o_h(1) \to 0$ should be uniform over $h \in H$. Moreover, the assertion (3.9) is not generally valid. A simple counter-example is (in notations of their paper):

$$Y_t = x_t\beta + u_t, \quad t = 1,2,\ldots, \ (\beta\text{: one-dimensional})$$

$$x_1 = 1/\sqrt{2}, \quad x_2 = 1 + \sqrt{2}/10, \quad x_3 = x_4 = \ldots = 1$$

$$u_1, u_2, \ldots \text{ i.i.d.}, \quad u_1 \sim N(0,1).$$

It is easy to verify that all conditions, including the uniqueness assumption and nonlattice condition, are satisfied. But it can easily be shown that

$$P_T\Big(Z_T(s,h) \in C[0,1]\Big) = 0$$

for h = 1 (which belongs to $H = \{1, 2, 3, \ldots T\}$) when $T = 2,4,6,\ldots$, and (3.9) breaks down (see Appendix 1).

Bloomfield and Steiger (1983) advanced a proof of (1.4) under the assumption that $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots$ are observations of a random vector $\underset{\sim}{X}$ with a positive

definite covariance matrix, and $(\underset{\sim}{x}_1, Y_1)$, $(\underset{\sim}{x}_2, Y_2)$, ... are stationary and ergodic. Unfortunately they failed to notice that for $\{g_n(\underline{c})\}$ (defined by (6) on p.45 of their book) to be equicontinuous, $g_n(c)$ must be defined as $\sum_{i=1}^n h_n\left(r_i(\underline{c})\right)/n$, and not $\sum_{i=1}^n h_n\left(r_i(\underline{c})\right)$ as in their book. But if $g_n(\underline{c})$ is defined as $\sum_{i=1}^n h_n\left(r_i(\underline{c})\right)/n$, the assertion $n^{-1}[\underset{\sim}{D}_n(\underset{\sim}{c}_n)] \to 2f(0)\underset{\sim}{C}$ on p.47 should be $[\underset{\sim}{D}_n(\underset{\sim}{c}_n)] \to 2f(0)\underset{\sim}{C}$, and one can only obtain $\underset{\sim}{c}_n - \underset{\sim}{a}_n \to 0$ in probability, not the crucial assertion (8) on p.46 of their book, and the proof breaks down. Besides, they made the mistake that the function $h_n(t)$ defined on p.45 of their book has no second order derivative at $t = \pm n^{-p}$, making the relation (12) on p.47 invalid.

Meanwhile Amemiya (1982) gave a proof of (1.4) by approximating the absolute value function with a twice-differentiable one. He made, in addition to (1.5), the assumption that $\{\underset{\sim}{x}_i\}$ is a bounded sequence, and that $\underset{\sim}{\beta}_0$ is confined in a compact region. Unfortunately his proof, too, is invalid. One problem is that his assertions (in notations of his paper) $A_1 \to 0$ in (3.12) and $B_1 \to 0$ $\left(\text{in } (3.22)\right)$ are both incorrect. Quite contrary, we have shown by simple arguments that actually $A_1 \to \infty$, $B_1 \to \infty$ in probability (see Appendix 2). Another crucial point is that in order to show $T(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*)'(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*) \to 0$ in probability as $T \to \infty$, (3.11) should be understood as $\sup\{|S^*(\underset{\sim}{\beta}) - S(\underset{\sim}{\beta})| : \underset{\sim}{\beta} \in D\} \to 0$ in probability (D is the parameter space), while his argument, even freed from the error indicated above, is obviously not sufficient for this.

When the regression model contains a constant term: $Y_i = \alpha_0 + \underset{\sim}{x}_i'\underset{\sim}{\beta}_0 + e_i$, $i = 1,2,\ldots$, Bloomfield and Steiger (1983, p.62 Lemma 1) noticed the interesting fact that the ML_1N estimator $\hat{\underset{\sim}{\beta}}_n$ of $\underset{\sim}{\beta}_0$ is in fact a special case of a class of rank estimators introduced by Jaeckel (1972). Jackel showed that his

estimator is asymptotically equivalent to an estimator introduced by Jurečková (1971). From this a proof of asymptotic normality of the ML_1N estimator $\hat{\beta}_n$ can be obtained by using the theorem proved by Jurečková (1971). However, this does not give a satisfactory solution of the problem for the following two reasons: First, Jurečková's theorem imposes very cumbersome conditions on the sequence $\{x_i\}$ which are difficult to verify. Her theorem also requires the existence of Fisher information of the density f of the error, so f must be positive and absolute continuous on R'. Even the simple uniform distribution $R(-1,1)$ does not meet this condition. Second, the theorem so obtained cannot deal with the case of (1.1) in which no constant term is present. If such a constant is present, the theorem cannot deal with this term.

Dupačová (1987) proved a theorem concerning the asymptotic normality of possible-constrained ML_1N estimates in case that $\{x_i\}$ is a random sequence. Her theorem, when applied to the unconstrained case, gives roughly the result stated by Bloomfield and Steiger (1983), as mentioned earlier. There is a mathematically undesirable condition in her theorem: $\|x_i\|$ possesses a finite moment of third order.

It is the purpose of this paper to give a rigorous proof of (1.4) under minimum conditions. First, in the i.i.d. case, we have the following theorem:

THEOREM 1. Suppose that in model (1.1), $e_1, e_2, \ldots$ are independent and identically distributed with a common distribution function F, and the following two conditions are satisfied.

1. There exists $\Delta > 0$ such that $f(u) = F'(u)$ exists when $|u| \leq \Delta$, f is continuous at 0, $f(0) > 0$ and $F(0) = 1/2$.

2. $\underset{\sim}{S}_n$ is nonsingular for some n, and

$$\lim_{n\to\infty} \max_{1\le i\le n} \underset{\sim}{x}_i' \underset{\sim}{S}_n^{-1} \underset{\sim}{x}_i = 0. \tag{1.6}$$

Then (1.4) is true.

Remark. The condition (1.6) is exactly the same as that which guarantees the asymptotic normality of the Least Squares estimate of $\underset{\sim}{\beta}$ in case that $\{e_i\}$ is i.i.d. and $Ee_1 = 0$, $0 < Ee_1^2 < \infty$. It was expected that the conditions ensuring the asymptotic normality of Minimum L_1-Norm estimate might be more stringent (as compared with the LS case), as the Minimum L_1-Norm estimate is nonlinear while the LS estimate is linear.

COROLLARY 1. If $\{e_i\}$ is i.i.d., condition 1 is satisfied, and there exists constant sequence $\{g_n\}$ such that $g_n \to \infty$, $g_{n+1}/g_n \to 1$, and

$$\underset{\sim}{S}_n/g_n \to \underset{\sim}{A} \quad \text{positive definite.} \tag{1.7}$$

Then (1.4) is true.

Wu (1981) mentioned this condition in connection with the problem of consistency of LS estimates.

This corollary contains, as a special case, the result stated in Bassett and Koenker (1978). In turn it implies the following result:

COROLLARY 2. Suppose that $\{e_i\}$ is i.i.d., condition 1 is satisfied, and $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots$ are i.i.d. observations of a random vector $\underset{\sim}{X}$ such that $E(\underset{\sim}{X}\underset{\sim}{X}')$ is positive definite, $\{x_i\}$ and $\{e_i\}$ are independent. Then with probability one (for almost every sample sequence $\{x_i\}$), (1.4) is true.

Of course, we need not consider (1.4) as a conditional statement: it is also true unconditionally. Thus we reach the conclusion stated in Bloomfield and Steiger (1983).

COROLLARY 3. The Minimum L_1-Norm estimate $\hat{\beta}_n$ is weak consistent under the conditions of Theorem 1 (see the remark after Lemma 4).

This assertion follows from (1.4), and the fact that

$$S_n^{-1} \to 0. \tag{1.8}$$

For a proof of (1.8), fix m such that S_m is positive definite. For any positive integer N, denote by $\rho_{N1} \leq \rho_{N2} \leq \cdots \leq \rho_{Np}$ the eigenvalues of S_N. Then by a result of von Neumann (1937), we have

$$\operatorname{tr}(S_m S_n^{-1}) \geq \sum_{i=1}^{p} \rho_{mi}/\rho_{ni}, \quad m \leq n. \tag{1.9}$$

But, by (1.6)

$$\operatorname{tr}(S_m S_n^{-1}) = \sum_{j=1}^{m} \operatorname{tr}(x_j x_j' S_n^{-1}) = \sum_{j=1}^{m} \operatorname{tr}(x_j' S_n^{-1} x_j)$$

$$= \sum_{j=1}^{m} x_j' S_n^{-1} x_j \leq m \max_{1 \leq i \leq m} x_i' S_n^{-1} x_i \to 0, \quad \text{as } n \to \infty. \tag{1.10}$$

Since $\rho_{m1} > 0$, from (1.9) and (1.10), we have $\lim_{n\to\infty} \rho_{n1} = \infty$, and (1.8) is proved.

From Corollary 3 it follows that if we use a subset G containing β_0 as an inner point to replace R^p in (1.2), and denote the resulting solution by $\hat{\beta}_n(G)$, we shall have $P(\hat{\beta}_n(G) \neq \hat{\beta}_n) \to 0$ as $n \to \infty$. Hence (1.4) is still true if $\hat{\beta}_n$ is replaced by $\hat{\beta}_n(G)$.

In passing we note that Y. Wu (1987) proved the strong consistency of $\hat{\beta}_n$ under conditions slightly stronger than those of Theorem 1. It does not seem possible to give a proof under the conditions of Theorem 1.

In practical applications there is usually a constant term in the re-

gression function, and instead of (1.1) we have the form

$$Y_i = \alpha_0 + \underset{\sim}{x}_i'\underset{\sim}{\beta}_0 + e_i, \quad i = 1,\ldots,n,\ldots . \tag{1.11}$$

Although (1.11), as a special case of (1.1), can be dealt with by Theorem 1, for inference purpose it will be convenient to have a theorem formulated in the following manner.

THEOREM 2. Write $(\hat{\alpha}_n, \hat{\underset{\sim}{\beta}}_n')$ the Minimum L_1-Norm estimate of $(\alpha_0, \underset{\sim}{\beta}_0')$, and

$$\bar{\underset{\sim}{x}}_n = (\underset{\sim}{x}_1 + \ldots + \underset{\sim}{x}_n)/n, \qquad \underset{\sim}{T}_n = \sum_{i=1}^{n} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_n)(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_n)'. \tag{1.12}$$

Suppose that $\{e_i\}$ is i.i.d., condition 1 of Theorem 1 is satisfied, $\underset{\sim}{T}_n$ is nonsingular for some n, and that

$$\lim_{n\to\infty} \max_{1\le i\le n} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_n)'\underset{\sim}{T}_n^{-1}(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}_n) = 0. \tag{1.13}$$

Then as $n \to \infty$, we have

$$2f(0)\underset{\sim}{T}_n^{1/2}(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}_0) \xrightarrow{\mathcal{L}} N(\underset{\sim}{0}, \underset{\sim}{I}_p) \tag{1.14}$$

$$\frac{2f(0)\sqrt{n}}{\sqrt{1 + n\bar{\underset{\sim}{x}}_n'\underset{\sim}{T}_n^{-1}\bar{\underset{\sim}{x}}_n}}(\hat{\alpha}_n - \alpha_0) \xrightarrow{\mathcal{L}} N(0,1). \tag{1.15}$$

Also, the two variables $2f(0)\underset{\sim}{T}_n^{1/2}(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}_0)$ and $2f(0)\sqrt{n}\{(\hat{\alpha}_n - \alpha_0) + \bar{\underset{\sim}{x}}_n'(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}_0)\}$ are asymptotically independent.

We note that the weak consistency of $\hat{\alpha}_n$ and $\hat{\underset{\sim}{\beta}}_n$ still holds true. For $\hat{\underset{\sim}{\beta}}_n$ the assertion follows from (1.14) and $\underset{\sim}{T}_n^{-1} \to 0$, which is a consequence of (1.13), in much the same way as (1.8) is a consequence of (1.6). For $\hat{\alpha}_n$ the assertion follows from (1.15) and $\bar{\underset{\sim}{x}}_n'\underset{\sim}{T}_n^{-1}\underset{\sim}{x}_n \to 0$, which is a trivial consequence of (1.13) and the fact that $\underset{\sim}{T}_n^{-1} \to 0$.

Corollary 2 can trivially be modified to accomodate the model (1.11). All we need to do is to replace the matrix $E(\underset{\sim}{X}\underset{\sim}{X}')$ by the covariance matrix $E(\underset{\sim}{X} - E\underset{\sim}{X})(\underset{\sim}{X} - E\underset{\sim}{X})'$.

The above two theorems can easily be extended to the case that e_1, e_2, ... are independent but not necessarily identically distributed. We shall only give a formulation of the following result.

THEOREM 3. Suppose that in model (1.1), $e_1, e_2, \ldots$ are independent, the distribution function $F_i(x)$ is differentiable over an interval $(-\Delta, \Delta)$, $F_i(0) = 1/2$, $i = 1,2,\ldots$ and $\Delta > 0$ does not depend on i. Write $f_i(x) = F_i'(x)$. Suppose that $\{f_i(x)\}$ is equicontinuous at $x = 0$ and $0 < \inf_i f_i(0) \leq \sup_i f_i(0) < \infty$. Finally, suppose that (1.6) is true. Then as $n \to \infty$, we have

$$2\underset{\sim}{S}_n^{-1/2} \sum_{i=1}^{n} f_i(0)\underset{\sim}{x}_i\underset{\sim}{x}_i'(\hat{\underset{\sim}{\beta}}_n - \underset{\sim}{\beta}_0) \xrightarrow{\mathcal{L}} N(\underset{\sim}{0}, \underset{\sim}{I}_p). \tag{1.16}$$

Our main task is to prove Theorem 1. Once this is achieved, only some trivial modifications are needed in proving Theorem 3, and much the same can be said about Theorem 2. To prove Theorem 1, it will be found convenient to reformulate the original problem in the following manner. Write

$$\underset{\sim}{S}_n^{-1/2}\underset{\sim}{x}_i = \underset{\sim}{x}_{ni}, \quad i = 1,\ldots,n, \qquad \underset{\sim}{S}_n^{1/2}\underset{\sim}{\beta}_0 = \underset{\sim}{\beta}_{n0}, \qquad Y_{ni} = Y_i, \qquad e_{ni} = e_i.$$

Then (1.1) has the form

$$Y_{ni} = \underset{\sim}{x}_{ni}'\underset{\sim}{\beta}_{n0} + e_{ni}, \quad i = 1,\ldots,n, \quad n = 1,2,\ldots \tag{1.17}$$

with

$$\sum_{i=1}^{n} \underset{\sim}{x}_{ni}\underset{\sim}{x}_{ni}' = \underset{\sim}{I}_p, \quad n = 1,2,\ldots \tag{1.18}$$

Denote by $\hat{\beta}_n$ the Minimum L_1-Norm estimate of β_{n0} in (1.17). Then we have the following general theorem, which includes Theorem 1 as a special case.

THEOREM 1'. Suppose that in model (1.17), $x_{n1}, \ldots, x_{nn}$ are known p-vectors satisfying (1.18), $e_{n1}, \ldots, e_{nn}$ are i.i.d. variables whose common distribution function F does not depend on n and satisfies the condition 1 of Theorem 1. Also, assume that

$$d_n \equiv \max_{1 \leq i \leq n} \|x_{ni}\| \to 0, \quad \text{as } n \to \infty \tag{1.19}$$

where $\|\cdot\|$ is the Euclidean norm in R^p. Then as $n \to \infty$ we have

$$2f(0)(\hat{\beta}_n - \beta_{n0}) \xrightarrow{\mathcal{L}} N(0, I_p). \tag{1.20}$$

This theorem will be proved in Section 3. Part of the reasoning is contained in Section 2 in the form of several preliminary lemmas.

2. SOME LEMMAS

LEMMA 1 (Bennett). Suppose that $\xi_1, \ldots, \xi_n$ are independent, $E\xi_i = 0$, $|\xi_i| \leq b < \infty$, $i = 1,\ldots,n$, where b is a constant. Write $V = \sum_{i=1}^n \mathrm{Var}(\xi_i)/n$. Then for any $\varepsilon > 0$, we have

$$P\left(\left|\sum_{i=1}^n \xi_i/n\right| \geq \varepsilon\right) \leq 2\exp\left(-n\varepsilon^2/(2V + 2b\varepsilon)\right). \tag{2.1}$$

For a proof, see Bennett (1962).

In model (1.17), we can assume that

$$\beta_{n0} = 0 \tag{2.2}$$

without any loss of generality. This we shall always do in the sequel, and we have $Y_{ni} = e_{ni}$. For any vector $a = (a_1, \ldots, a_p)'$ in R^p, write $|a| = \max_{1\leq i\leq p}|a_i|$. $I(A)$ will be used to denote the indicator of the set A.

LEMMA 2. Suppose that in model (1.17) the conditions of Theorem 1' are satisfied, with the possible exception of (1.19). Then we have

$$I\left(\left|\sum_{i=1}^n \mathrm{sgn}(Y_{ni} - x_{ni}'\hat{\beta}_n)x_{ni}\right| \geq (p+1)d_n\right) \leq I\left(|\hat{\beta}_n| \geq \Delta/(\sqrt{p}d_n)\right) \tag{2.3}$$

with probability one, where $\mathrm{sgn}(0) = 0$, $\mathrm{sgn}(a) = a/|a|$ for $a \neq 0$.

Proof. Since by definition $\hat{\beta}_n$ is a minimization point of $\sum_{i=1}^n |Y_{ni} - x_{ni}'\beta|$ as a function of $\beta \in R^p$, for any unit vector $\theta \in R^p$, we have by taking directional derivative

$$-\sum_{i=1}^n \mathrm{sgn}(Y_{ni} - x_{ni}'\hat{\beta}_n)x_{ni}'\theta I(Y_{ni} \neq x_{ni}'\hat{\beta}_n) + \sum_{i=1}^n |x_{ni}'\theta| I(Y_{ni} = x_{ni}'\hat{\beta}_n) \geq 0.$$

This implies, in view of the arbitrariness of θ, that

$$\left|\sum_{i=1}^{n} \operatorname{sgn}(Y_{ni} - \underset{\sim}{x}'_{nk}\hat{\underset{\sim}{\beta}}_n)\underset{\sim}{x}_{ni}\right| \leq \sum_{i=1}^{n} |\underset{\sim}{x}_{ni}| I(Y_{ni} = \underset{\sim}{x}'_{ni}\hat{\underset{\sim}{\beta}}_n). \tag{2.4}$$

Now suppose that $|\hat{\underset{\sim}{\beta}}_n| < \Delta/(\sqrt{p}\, d_n)$. If $1 \leq i_1 \leq i_2 < \ldots < i_{p+1} \leq n$ and

$$Y_{nk} = \underset{\sim}{x}'_{nk}\hat{\underset{\sim}{\beta}}_n, \quad k = i_1, \ldots, i_{p+1}.$$

Then find real constants $c_1, \ldots, c_{p+1}$ not equaling to zero simultaneously such that $c_1 x_{ni_1} + \ldots + c_{p+1} x_{ni_{p+1}} = 0$. We have

$$c_1 Y_{ni_1} + \ldots + c_{p+1} Y_{ni_{p+1}} = 0. \tag{2.5}$$

Considering this, and the fact that $|\hat{\underset{\sim}{\beta}}_n| < \Delta/(\sqrt{p}\, d_n)$ implies $|\underset{\sim}{x}'_{ni}\hat{\underset{\sim}{\beta}}_n| \leq \Delta$, we reach the following conclusion:

The event $\left\{\sum_{i=1}^{n} |x_{ni}| I(Y_{ni} = \underset{\sim}{x}'_{ni}\hat{\underset{\sim}{\beta}}_n) \geq (p+1)d_n, \text{ and } |\hat{\underset{\sim}{\beta}}_n| < \Delta/(\sqrt{p}\, d_n)\right\}$

$\subset$ the event $\bigcup_{1 \leq i_1 < \ldots < i_{p+1} \leq n} \Big\{$there exists constants $c_1, \ldots, c_{p+1}$

depending only on $x_{ni_1}, \ldots, x_{ni_{p+1}}$, not all zero, such that

$$\sum_{j=1}^{p+1} c_j Y_{ni_j} = 0, \text{ and } |Y_{ni_j}| < \Delta,\ j = 1, \ldots, p+1\Big\}. \tag{2.6}$$

Since $Y_{ni} = e_{ni}$, $e_{n1}, \ldots, e_{nn}$ are i.i.d. variables whose common distribution function is continuous over $(-\Delta, \Delta)$, it is seen that the probability of the event on the right-hand side of (2.6) is zero. This fact, taken together with (2.4), implies (2.3), and the lemma is proved.

In order to introduce the crucial Lemma 3, we have to define some notations.

By (1.19) it follows that there exists constant sequence $\{\mu_n\}$ of positive

even integers such that as $n \to \infty$,

$$\mu_n \to \infty, \quad \mu_n^2 d_n \to 0, \quad \sqrt{p}\,\mu_n^2 d_n < 1/2, \quad \text{for } n \text{ large.} \tag{2.7}$$

In the following we assume that the conditions of Theorem 1' are met. Write

$$M \equiv M_n \equiv [\log n/\log \mu_n], \quad r_m \equiv \mu_n^m, \quad m = 1,\ldots,M, \quad r_{M+1} = n. \tag{2.8}$$

$$D \equiv D_n \equiv \{\beta = (\beta_1, \ldots, \beta_p)' : -\mu_n \le \beta_i < \mu_n,\ i = 1,\ldots,p\}. \tag{2.9}$$

Note that $M \ge 1$ for n large. This is true because by (1.18) we have $nd_n^2 \ge 1$, therefore $\mu_n < n^{1/4}$. Partition D into a number of intervals $\tilde{D}_1$, $\tilde{D}_2, \ldots, \tilde{D}_{J_1}$ each having the form $\{\underset{\sim}{x} = (x^{(1)}, \ldots, x^{(p)})' : a_i \le x^{(i)} < b_i,$ $i = 1,\ldots,p\}$, such that

$$J_1 \le \mu_n^{2p}, \quad L(\tilde{D}_i) \le 1, \quad \tilde{D}_i \cap \tilde{D}_j = \emptyset, \quad i,j = 1,\ldots,J_1, \quad i \ne j$$

where $L(A)$ is defined as $\sup\{|\underset{\sim}{u} - \underset{\sim}{v}| : \underset{\sim}{u} \in A, \underset{\sim}{v} \in A\}$. Now each subset $\tilde{D}_k$ is again partitioned into a number of disjoint intervals $\tilde{D}_{k1}, \tilde{D}_{k2}, \ldots, \tilde{D}_{kJ_2}$ which can further be partioned. This process is defined inductively as follows: Suppose that after the (m-1)-th round we have partitioned D into $\{\tilde{D}_{j_1 \cdots j_{m-1}}\}$, then in the m-th round we take each $\tilde{D}_{j_1 \cdots j_{m-1}}$ and partition it into a number of disjoint intervals $\{\tilde{D}_{j_1 \cdots j_{m-1}\ell} : \ell = 1,2,\ldots,J_m\}$, such that

$$J_m \le \mu_n^{2p}, \quad L(\tilde{D}_{j_1 \cdots j_{m+1}\ell}) \le \mu_n^{-2(m-1)}, \quad \ell = 1,\ldots,J_m. \tag{2.10}$$

The process ends with the completion of the (M+1)-th round. Denote

$$G_m = \{\tilde{D}_{j_1 \cdots j_m}\} = \text{the partitioning of } D \text{ after the m-th round}$$
$$m = 1,2,\ldots,M+1. \tag{2.11}$$

A typical element of G_m is denoted by B or $\tilde{D}_{j_1 \ldots j_m}$, and a chosen point in it is denoted by $\underset{\sim}{b}$ or $\underset{\sim}{b}_{j_1 \ldots j_m}$. Put

$$\psi_{ni}(\beta) = \mathrm{sgn}(e_{ni} - \underset{\sim}{x}'_{ni}\underset{\sim}{\beta}) - \mathrm{sgn}(e_{ni}) \tag{2.12}$$

$$\lambda_{ni}(\beta) = 2(F(\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}) - 1/2) = -E\psi_{ni}(\underset{\sim}{\beta}) \tag{2.13}$$

$$t_{ni}(\underset{\sim}{\beta}) = \psi_{ni}(\underset{\sim}{\beta}) + \lambda_{ni}(\underset{\sim}{\beta}). \tag{2.14}$$

By (2.7), when $\underset{\sim}{\beta} \in D$ and $i \leq n$, we have, for large n,

$$|\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}| \leq \sqrt{p}\, d_n |\underset{\sim}{\beta}| < (2\mu_n)^{-1} \tag{2.15}$$

and for $\underset{\sim}{b} \in B$, $\underset{\sim}{b}^* \in B$ where $B \in G_m$, we have, for large n,

$$|\underset{\sim}{x}'_{ni}(\underset{\sim}{b}-\underset{\sim}{b}^*)| \leq \sqrt{p}\, d_n |\underset{\sim}{b}-\underset{\sim}{b}_*| \leq (2\mu_n^{2m})^{-1}, \quad 1 \leq i \leq n. \tag{2.16}$$

From (2.15) and the conditions imposed on the distribution function F (see condition 1 of Theorem 1), it follows that we can find a constant $C > 1$ such that for $\underset{\sim}{\beta} \in D$ and $1 \leq i \leq n$:

$$2\{F(\underset{\sim}{x}'_{ni}\underset{\sim}{\beta} + \mu_n^{-2m}) - F(\underset{\sim}{x}'_{ni}\underset{\sim}{\beta} - \mu_n^{-2m})\} \leq C\mu_n^{-2m} \tag{2.17}$$

$$|\lambda_{ni}(\underset{\sim}{\beta})| \leq C\mu_n^{-1} \tag{2.18}$$

$$P(|e_{ni}| \leq \sqrt{p}\, d_n \mu_n) \leq C d_n \mu_n. \tag{2.19}$$

Let $\{a_{ni},\ 1 \leq i \leq n,\ n = 1,2,\ldots\}$ be a triangular array of real numbers, satisfying

$$\sum_{i=1}^{n} a_{ni}^2 = 1, \quad |a_{ni}| \leq \mu_n^{-m/2} \quad \text{for} \quad i \geq r_m + 1 \tag{2.20}$$

$$n = 1,2,\ldots, \quad m = 1,2,\ldots,M+1.$$

Define the following quantities:

$$Q_m = \sum_{B\in G_m} P\left\{\sup_{\underset{\sim}{\beta}\in B}\left|\sum_{i=r_m+1}^{n} a_{ni}\left(t_{ni}(\underset{\sim}{\beta}) - t_{ni}(\underset{\sim}{b})\right)\right| \geq \varepsilon 3^{-m}\right\},$$

$$1 \leq m \leq M+1, \tag{2.21}$$

$$U_m = \sum_{B\in G_m} P\left\{\sup_{\underset{\sim}{\beta}\in B}\left|\sum_{i=r_m+1}^{r_{m+1}} a_{ni}\left(t_{ni}(\underset{\sim}{\beta}) - t_{ni}(\underset{\sim}{b})\right)\right| \geq \varepsilon 3^{-(m+1)}\right\} \equiv \sum_{b\in G_m} U(m,B),$$

$$1 \leq m \leq M, \tag{2.22}$$

$$V_m = \sum_{j_1,\ldots,j_{m+1}} P\left\{\left|\sum_{i=r_{m+1}+1}^{n} a_{ni}\left(t_{ni}(\underset{\sim}{b}_{j_1\ldots j_m}) - t_{ni}(\underset{\sim}{b}_{j_1\ldots j_m j_{m+1}})\right)\right| \geq \varepsilon 3^{-(m+1)}\right\}$$

$$\equiv \sum_{j_1,\ldots,j_{m+1}} V(m,j_1,\ldots,j_{m+1}), \qquad 1 \leq m \leq M, \quad V_M = 0. \tag{2.23}$$

In the above expressions $\varepsilon \in (0,1)$, r_m, $m = 1,\ldots,M,M+1$, has been defined in (2.8). Further, in the definition of V_m, the summation runs over all such $(j_1, \ldots, j_{m+1})$ that $\tilde{D}_{j_1\ldots j_{m+1}}$ is a member of G_{m+1}, and $\underset{\sim}{b}_{j_1\ldots j_t}$ is understood as the point chosen in $\tilde{D}_{j_1\ldots j_t} \in G_t$, while in the definition of Q_m and U_m, b is the point chosen in a member B of G_m, as stated earlier. By these definitions it is easily seen that

$$Q_m \leq U_m + V_m + Q_{m+1}, \quad m = 1,\ldots,M. \tag{2.24}$$

Therefore, on noticing that $Q_{M+1} = 0$, we have

$$Q_1 \leq \sum_{m=1}^{M} (U_m + V_m). \tag{2.25}$$

LEMMA 3. Suppose that the conditions of Theorem 1' are met, and $\{a_{ni}\}$ satisfies (2.20), $\varepsilon \in (0,1)$. Then when n is large we have

$$U_m + V_m \leq 4\mu_n^{2pm}\exp\{-48^{-1}\varepsilon(9^{-1}\mu_n)^{m/2}\}, \quad m = 1,\ldots,M \tag{2.26}$$

where C is the constant appearing in (2.17)-(2.19).

Proof. Define

$$\eta_{ni} \equiv \eta_{ni}(B) = \begin{cases} 2, & \text{when } |e_{ni} - \underset{\sim}{x}'_{ni}\underset{\sim}{b}| < \mu_n^{-2m} \\ 0, & \text{otherwise} \end{cases} \tag{2.27}$$

where $B \in G_m$ and $\underset{\sim}{b}$ is the point chosen in B. By (2.14), (2.16) and (2.17), we have

$$\sup_{\beta \in B}\left|\sum_{i=r_m+1}^{r_{m+1}} a_{ni}\left(t_{ni}(\underset{\sim}{\beta}) - t_{ni}(\underset{\sim}{b})\right)\right|$$

$$\leq \sum_{i=r_m+1}^{r_{m+1}} |a_{ni}|(\eta_{ni} - E\eta_{ni}) + 2\sum_{i=r_m+1}^{r_{m+1}} |a_{ni}|E\eta_{ni}. \tag{2.28}$$

From (2.16), (2.17) and (2.27), we have

$$E\eta_{ni} \leq C\mu_n^{-2m}, \quad i = r_m+1,\ldots,r_{m+1}. \tag{2.29}$$

By (2.20), we have for n large

$$2\sum_{i=r_m+1}^{r_{m+1}} |a_{ni}|E\eta_{ni} \leq 2\Big(\sum_{i=r_m+1}^{r_{m+1}} (E\eta_{ni})^2\Big)^{1/2} \leq 2(\mu_n^{m+1}C^2\mu_n^{-4m})^{1/2}$$

$$\leq 2C\mu_n^{-m} < \varepsilon 2^{-1}3^{-(m+1)}. \tag{2.30}$$

From (2.1), (2.28), (2.30) and by Lemma 1, we obtain

$$U(m,B) \le P\left\{\left|\sum_{i=r_m+1}^{r_{m+1}} |a_{ni}|(\eta_{ni} - E\eta_{ni})\right| \ge \varepsilon 2^{-1} 3^{-(m+1)}\right\}$$

$$\le 2\exp\left\{-\varepsilon^2 2^{-2} 3^{-(2m+2)} \Big/ \left[\left(\max_{r_m < i \le r_{m+1}} E\eta_{ni}^2 + \varepsilon 3^{-(m+1)} \max_{i > r_m} |a_{ni}|\right)\right]\right\} \tag{2.31}$$

$U(m,B)$ is defined in (2.22). From (2.17), (2.20) and (2.29), we have, for large n,

$$\max_{i > r_m} |a_{ni}| \le \mu_n^{-m/2},$$

$$\max_{r_m < i \le r_{m+1}} E\eta_{ni}^2 \le 2C\mu_n^{-2m} \le \mu_n^{-m/2} \varepsilon 3^{-(m+1)}.$$

Therefore

$$U(m,B) \le 2\exp(-\varepsilon 3^{-(m+1)} \mu_n^{m/2}/16) \tag{2.32}$$

which implies

$$U_m \le 2\mu_n^{2pm} \exp(-\varepsilon 3^{-(m+1)} \mu_n^{m/2}/16). \tag{2.33}$$

By (2.20),

$$4 \max_{i > r_{m+1}} |a_{ni}| \varepsilon 3^{-(m+1)} \le 4\varepsilon 3^{-(m+1)} \mu_n^{-(m+1)/2}. \tag{2.34}$$

By (2.16) and (2.17), we have, on denoting $g_{ni} = t_{ni}(\underset{\sim}{b}_{j_1 \ldots j_m}) - t_{ni}(\underset{\sim}{b}_{j_1 \ldots j_{m+1}})$, that

$$Eg_{ni}^2 \le 4\left|F(\underset{\sim}{x}'_{ni}\underset{\sim}{b}_{j_1 \ldots j_m}) - F(\underset{\sim}{x}'_{ni}\underset{\sim}{b}_{j_1 \ldots j_{m+1}})\right|$$

$$\le 2C\mu_n^{-2m} \le 4\varepsilon 3^{-(m+1)} \mu_n^{-(m+1)/2}. \tag{2.35}$$

Here $\underset{\sim}{b}_{j_1 \ldots j_m}$, $\underset{\sim}{b}_{j_1 \ldots j_{m+1}}$ were explained following the definition of V_m.

From (2.23), (2.34), (2.35), and Bennett inequality (2.1), we have

$$V(m, j_1, \ldots, j_m) \le 2\exp\left\{-\varepsilon^2 3^{-(2m+2)} \Big/ \left[2\left(\max_{i<r_{m+1}} Eg_{ni}^2 + 4 \max_{i>r_{m+1}} |a_{ni}|\varepsilon 3^{-(m+1)}\right)\right]\right\}$$
$$\le 2\exp\left\{-16^{-1}\varepsilon(\mu_n/9)^{(m+1)/2}\right\}$$

which in turn implies

$$V_m \le 2\mu_n^{2p(m+1)}\exp\{-16^{-1}\varepsilon(\mu_n/9)^{(m+1)/2}\} \le 2\mu_n^{2pm}\exp\{-16^{-1}\varepsilon(\mu_n/9)^{m/2}\}. \quad (2.36)$$

Finally, for n large, (2.26) follows from (2.33) and (2.36). Lemma 3 is proved.

LEMMA 4. Under the assumptions of Theorem 1', we have

$$\lim_{n\to\infty} P(|\hat{\underset{\sim}{\beta}}_n| > v_n) = 0. \quad (2.37)$$

for any constant sequence $\{v_n\}$ such that $\lim_{n\to\infty} v_n = \infty$. (Here we assume (2.2).)

Proof. Without losing generality we may assume $d_n v_n^2 \le 1$. Define

$$\tilde{D} \equiv \tilde{D}_n \equiv \{\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_p)': -v_n \le \beta_i \le v_n, \ i = 1, \ldots, p\}$$

$$\Phi_{ni}(\underset{\sim}{\beta}) = |e_{ni}| - |e_{ni} - \underset{\sim}{x}'_{ni}\underset{\sim}{\beta}|$$

$$\Lambda_{ni}(\underset{\sim}{\beta}) = E\left(\Phi_{ni}(\underset{\sim}{\beta})\right)$$

$$R_{ni}(\underset{\sim}{\beta}) = \Phi_{ni}(\underset{\sim}{\beta}) - \Lambda_{ni}(\underset{\sim}{\beta}), \quad i = 1, \ldots, n; \quad n = 1, 2, \ldots .$$

The first step is to verify that

$$\lim_{n\to\infty} v_n^{-2} \sup_{\underset{\sim}{\beta}\in\tilde{D}} \left|\sum_{i=1}^{n} R_{ni}(\underset{\sim}{\beta})\right| = 0, \quad \text{in probability} \quad (2.38)$$

In order to do this, we partition the interval $\tilde{D}$ in exactly the same manner as we have previously done for D defined by (2.9), with μ_n replaced by v_n, and that instead of (2.8), we now define M as a positive integer satisfying

$$\sqrt{np}\, v_n^{-2M} < 2^{-1}\varepsilon 3^{-(M+1)} v_n^2 \tag{2.39}$$

where $\varepsilon > 0$ is an arbitrarily given constant. The existence of such M follows from the fact that $v_n \to \infty$. Also, the partitioning of $\tilde{D}$ after the m-th round will be denoted by $\tilde{G}_m$. A typical interval belonging to $\tilde{G}_m$ will be denoted by B or $B_{j_1,\ldots,j_m}$, and a point selected from it by $\underset{\sim}{b}$ or $\underset{\sim}{b}_{j_1,\ldots,j_m}$. Define

$$\tilde{V}_0 = \sum_{j_1} P\Big(\Big|\sum_{i=1}^n R_{ni}(\underset{\sim}{b}_{j_1})\Big| \geq \varepsilon v_n^2/3\Big)$$

$$\tilde{V}_m = \sum_{j_1,\ldots,j_{m+1}} P\Big(\Big|\sum_{i=1}^n \big(R_{ni}(b_{j_1\ldots j_m}) - R_{ni}(\underset{\sim}{b}_{j_1\ldots j_{m+1}})\big)\Big| \geq \varepsilon v_n^2/3^{m+1}\Big)$$

$$m = 1,2,\ldots,M;$$

$$\tilde{Q}_m = \sum_{j_1\ldots j_m} P\Big(\sup_{\underset{\sim}{\beta}\in B_{j_1\ldots j_m}} \Big|\sum_{i=1}^n \big(R_{ni}(\underset{\sim}{\beta}) - R_{ni}(\underset{\sim}{b}_{j_1\ldots j_m})\big)\Big| \geq \varepsilon v_n^2/3^m\Big)$$

$$m = 1,2,\ldots,M+1.$$

Note that for any $B \in \tilde{G}_{M+1}$ and $\underset{\sim}{\beta} \in B$ we have, in view of (2.39),

$$\Big|\sum_{i=1}^n \big(\Phi_{ni}(\underset{\sim}{\beta}) - \Phi_{ni}(\underset{\sim}{b})\big)\Big| \leq \sum_{i=1}^n |\underset{\sim}{x}'_{ni}(\underset{\sim}{\beta} - \underset{\sim}{b})| \leq \Big(n\sum_{i=1}^n \big(\underset{\sim}{x}'_{ni}(\underset{\sim}{\beta} - \underset{\sim}{b})\big)^2\Big)^{1/2}$$

$$= \sqrt{n}\,\|\underset{\sim}{\beta} - \underset{\sim}{b}\| \leq \sqrt{np}\,|\underset{\sim}{\beta} - \underset{\sim}{b}| \leq \sqrt{np}\, v_n^{-2M} < 2^{-1}\varepsilon 3^{-(M+1)} v_n^2.$$

Hence

$$\sup_{\underset{\sim}{\beta}\in B}\Big|\sum_{i=1}^n \big(R_{ni}(\underset{\sim}{\beta}) - R_{ni}(\underset{\sim}{b})\big)\Big| < \varepsilon 3^{-(M+1)} v_n^2$$

which implies that

$$\tilde{Q}_{M+1} = 0. \tag{2.40}$$

It is easy to see that

$$\tilde{Q}_m \le \tilde{V}_m + \tilde{Q}_{m+1}, \quad m = 1,\dots,M, \tag{2.41}$$

$$P\Big(v_n^{-2} \sup_{\underset{\sim}{\beta}\in\tilde{D}} \Big|\sum_{i=1}^{n} R_{ni}(\underset{\sim}{\beta})\Big| \ge \varepsilon\Big) \le \tilde{V}_0 + \tilde{Q}_1. \tag{2.42}$$

From (2.40)-(2.42), it follows that

$$P\Big(v_n^{-2} \sup_{\underset{\sim}{\beta}\in\tilde{D}} \Big|\sum_{i=1}^{n} R_{ni}(\underset{\sim}{\beta})\Big| \ge \varepsilon\Big) \le \sum_{m=0}^{M} \tilde{V}_m. \tag{2.43}$$

Since

$$|\Phi_{ni}(\underset{\sim}{b}_{j_1})| \le |\underset{\sim}{x}'_{ni}\underset{\sim}{b}_{j_1}| \le d_n v_n \le v_n^{-1}, \quad \text{since} \quad d_n v_n^2 \le 1.$$

$$\sum_{i=1}^{n}\Big(\mathrm{Var}\Big(\Phi_{ni}(\underset{\sim}{b}_{j_1}\Big)\Big) \le \sum_{i=1}^{n}(\underset{\sim}{x}'_{ni}\underset{\sim}{b}_{j_1})^2 = \|\underset{\sim}{b}_{j_1}\|^2 \le p v_n^2,$$

$$\begin{aligned}|\Phi_{ni}(\underset{\sim}{b}_{j_1\dots j_m}) - \Phi_{ni}(\underset{\sim}{b}_{j_1\dots j_{m+1}})| &\le |\underset{\sim}{x}'_{ni}(\underset{\sim}{b}_{j_1\dots j_m} - \underset{\sim}{b}_{j_1\dots j_{m+1}})| \\ &\le \|\underset{\sim}{x}_{ni}\|\cdot\|\underset{\sim}{b}_{j_1\dots j_m} - \underset{\sim}{b}_{j_1\dots j_{m+1}}\| \\ &\le p|\underset{\sim}{x}_{ni}|\cdot|\underset{\sim}{b}_{j_1\dots j_m} - \underset{\sim}{b}_{j_1\dots j_{m+1}}| \\ &\le p d_n v_n^{-2m+2} \le p v_n^{-2m},\end{aligned}$$

and

$$\sum_{i=1}^{n} \mathrm{Var}\left(\phi_{ni}(\underset{\sim}{b}_{j_1\ldots j_m}) - \phi_{ni}(\underset{\sim}{b}_{j_1\ldots j_{m+1}})\right) \le \sum_{i=1}^{n}\left(\underset{\sim}{x}'_{ni}(\underset{\sim}{b}_{j_1\ldots j_m} - \underset{\sim}{b}_{j_1\ldots j_{m+1}})\right)^2$$

$$= \|\underset{\sim}{b}_{j_1\ldots j_m} - \underset{\sim}{b}_{j_1\ldots j_{m+1}}\|^2$$

$$\le p|\underset{\sim}{b}_{j_1\ldots j_m} - \underset{\sim}{b}_{j_1\ldots j_{m+1}}|^2$$

$$\le p v_n^{-4m+4}.$$

Applying Lemma 1, we get

$$\tilde{V}_0 \le 2v_n^{2p}\exp\left(-3^{-2}\varepsilon^2 v_n^4/(2pv_n^2 + \varepsilon v_n)\right) \le 2v_n^{2p}\exp(-cv_n^2). \qquad (2.44)$$

$$\tilde{V}_m \le 2v_n^{2p(m+1)}\exp\{-(\varepsilon 3^{-m-1}v_n^2)^2/(2pv_n^{-4m+4} + \varepsilon 3^{-m-1}pv_n^{-2m+2})\}$$

$$\le 2v_n^{2p(m+1)}\exp(-c3^{-(m+1)}v_n^{2m+2}), \quad m \ge 1. \qquad (2.45)$$

Here $c > 0$ is a constant independent of n, m. From (2.43)-(2.45), we obtain

$$P\left(v_n^{-2}\sup_{\underset{\sim}{\beta}\in\tilde{D}}|\sum_{i=1}^{n} R_{ni}(\underset{\sim}{\beta})| \ge \varepsilon\right) \le 2\sum_{m=0}^{\infty} v_n^{2p(m+1)}\exp(-c3^{-(m+1)}v_n^{2m+2}). \qquad (2.46)$$

Since the right-hand side of (2.46) tends to zero as $n \to \infty$, we obtain (2.38).

Now we note that

$$\sup_{\underset{\sim}{\beta}\in\tilde{D}}\max_{1\le i\le n}|\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}| \le d_n v_n \le v_n^{-1} \to 0, \quad \text{as} \quad n \to \infty. \qquad (2.47)$$

Hence, considering condition 1 of Theorem 1, we have

$$E\Phi_{ni}(\underset{\sim}{\beta}) = \begin{cases} \int_0^{\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}} (2u - 2\underset{\sim}{x}'_{ni}\underset{\sim}{\beta})f(u)du, & \text{if} \quad \underset{\sim}{x}'_{ni}\underset{\sim}{\beta} \ge 0 \\ \int_{\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}}^{0} (2\underset{\sim}{x}'_{ni}\underset{\sim}{\beta} - 2u)f(u)du, & \text{if} \quad \underset{\sim}{x}'_{ni}\underset{\sim}{\beta} < 0 \end{cases}$$

$$= -f(0)(\underset{\sim}{x}'_{ni}\underset{\sim}{\beta})^2\left(1 + o(1)\right) \qquad (2.48)$$

for $\underset{\sim}{\beta} \in \tilde{D}$, $i = 1,\ldots,m$, n sufficiently large, where $o(1) \to 0$ as $n \to \infty$ uniformly for $\underset{\sim}{\beta} \in \tilde{D}$ and $1 \leq i \leq n$, in view of (2.47). From (2.38) and (2.48), we get

$$\lim_{n\to\infty} v_n^{-2} \sup_{\underset{\sim}{\beta}\in\tilde{D}} \left| \sum_{i=1}^{n} (|e_{ni}| - |e_{ni} - \underset{\sim}{x}_{ni}'\underset{\sim}{\beta}|) + f(0)\|\underset{\sim}{\beta}\|^2 \right| = 0, \quad \text{in probability}$$

which implies that

$$v_n^{-2}\Big(\sum_{i=1}^{n} |e_{ni}| - \inf_{|\underset{\sim}{\beta}|=v_n} \sum_{i=1}^{n} |e_{ni} - \underset{\sim}{x}_{ni}'\underset{\sim}{\beta}|\Big) \leq -f(0)\big(1 + o_p(1)\big) \tag{2.49}$$

where $o_p(1)$ tends to zero in probability as $n \to \infty$. Since $\sum_{i=1}^{n} |e_{ni} - \underset{\sim}{x}_{ni}'\underset{\sim}{\beta}|$ as a function of $\underset{\sim}{\beta}$ is convex over R^p, (2.49) implies

$$v_n^{-2}\Big(\sum_{i=1}^{n} |e_{ni}| - \inf_{|\underset{\sim}{\beta}|\geq v_n} \sum_{i=1}^{n} |e_{ni} - \underset{\sim}{x}_{ni}'\underset{\sim}{\beta}|\Big) \leq -f(0)\big(1 + o_p(1)\big). \tag{2.50}$$

Since $f(0) > 0$, (2.50) implies (2.37), and Lemma 4 is proved.

° *Remark.* Let us return temporarily to model (1.1) and consider the ML_1N estimate $\hat{\underset{\sim}{\beta}}_n$ defined by (1.2). It follows easily from Lemma 4 and (1.8) that $\hat{\underset{\sim}{\beta}}_n$ is a weakly consistent estimate of $\underset{\sim}{\beta}_0$. For convenience of presentation we formulated this fact as a corollary of Theorem 1. Now we see that the verification of this fact is, in fact, an important step in the proof of Theorem 1.

3. PROOF OF THE THEOREMS

Proof of Theorem 1. As mentioned earlier, we need only to prove Theorem 1'. We begin with verifying that as $n \to \infty$,

$$\sup\{\|\sum_{i=1}^{n} t_{ni}(\underset{\sim}{\beta})\underset{\sim}{x}_{ni}\|: \underset{\sim}{\beta} \in \tilde{D}\} \to 0, \quad \text{in probability.} \tag{3.1}$$

D and $t_{ni}(\underset{\sim}{\beta})$ were defined by (2.9) and (2.12)-(2.14), respectively. In view of (1.18), we need only to prove that as $n \to \infty$,

$$Q_0 \equiv P(\sup\{|\sum_{i=1}^{n} a_{ni}t_{ni}(\underset{\sim}{\beta})|: \underset{\sim}{\beta} \in D\} \geq \varepsilon) \to 0 \tag{3.2}$$

where $\{a_{ni}\}$ is an arbitrarily given constant array satisfying $\sum_{i=1}^{n} a_{ni}^2 = 1$, and $\varepsilon \in (0,1)$, also arbitrarily given.

Without loss of generality, assume $|a_{ni}| \geq \ldots \geq |a_{nn}|$. Choose μ_n and r_m according to (2.7) and (2.8), then (2.20) holds obviously. Define U_m, V_m, Q_m by (2.21)-(2.23) for $1 \leq m \leq M$, and

$$U_0 = P\{\sup_{\underset{\sim}{\beta}\in D}\left|\sum_{i=1}^{r_1} a_{ni}t_{ni}(\underset{\sim}{\beta})\right| \geq \varepsilon/3\} \tag{3.3}$$

$$V_0 =\{\sum_{B\in G_1} P\left|\sum_{i=r_1+1}^{n} a_{ni}t_{ni}(\underset{\sim}{b})\right| \geq \varepsilon/3\} \tag{3.4}$$

with $b \in B$. It is easily verified that with Q_0, U_0, V_0 defined by (3.2)-(3.4), (2.24) also holds true for $m = 0$. Hence

$$Q_0 \leq U_0 + V_0 + Q_1 \leq \sum_{m=0}^{M}(U_m + V_m). \tag{3.5}$$

Here M is defined by (2.8). Now we show that

$$U_0 + V_0 \to 0, \quad \text{as} \quad n \to \infty. \tag{3.6}$$

By (2.18), for $\underset{\sim}{\beta} \in D$, we have

$$\left|\sum_{i=1}^{n} a_{ni}\lambda_{ni}(\underset{\sim}{\beta})\right| \le \left(\sum_{i=1}^{n} \lambda_{ni}^2(\underset{\sim}{\beta})\right)^{1/2} \le \{\mu_n(C/\mu_n)^2\}^{1/2} \to 0, \quad \text{as} \quad n \to \infty \tag{3.7}$$

Further, $\psi_{ni}(\underset{\sim}{\beta}) = 0$ when $\underset{\sim}{\beta} \in D$ and $|e_{ni}| > \sqrt{p}\, d_n\mu_n (> |\underset{\sim}{x}'_{ni}\underset{\sim}{\beta}|)$, here $\psi_{ni}(\underset{\sim}{\beta})$ is defined by (2.12). From (2.7), (2.19) and (3.7), we obtain

$$U_0 \le P\{\sup_{\underset{\sim}{\beta} \in D}\left|\sum_{i=1}^{r_1} a_{ni}\psi_{ni}(\underset{\sim}{\beta})\right| \ne 0\} \le \sum_{i=1}^{r_1} P(|e_{ni}| \le \sqrt{p}\, d_n\mu_n) \le Cd_n\mu_n^2 \to 0. \tag{3.8}$$

Using (2.17)-(2.20) and employing the argument for proving (2.36), it can be shown that there exists constant $C_1 > 0$ such that

$$V_0 \le 2\mu_n^{2p}\exp(-C_1\mu_n^{1/2}) \to 0, \quad \text{as} \quad n \to \infty. \tag{3.9}$$

Now (3.6) follows from (3.8) and (3.9). Further, by Lemma 3, we have, for n large,

$$\sum_{m=1}^{M}(U_m + V_m) \le 4\sum_{m=1}^{\infty}\mu_n^{2pm}\exp\left(-48^{-1}\varepsilon(\mu_n/9)^{m/2}\right). \tag{3.10}$$

Since the function $x^{3pm}\exp\{-a(\frac{x}{b})^{m/2}\}$, $x > 0$, attains its maximum at $x = b(6p/a)^{2/m}$, we have

$$\mu_n^{2pm}\exp\{-16^{-1}\varepsilon(\mu_n/9)^{m/2}\} = \mu_n^{-pm}\cdot\mu_n^{3pm}\exp\{-16^{-1}\varepsilon(\mu_n/9)^{m/2}\}$$

$$\le \mu_n^{-pm}9^{3pm}(96p/\varepsilon)^{6p}e^{-6p} \le (\mu_n/729)^{pm}(96p/\varepsilon)^{6p}.$$

Hence, noting that $\mu_n \to \infty$, we obtain

$$\sum_{m=1}^{M}(U_m + V_m) \le (\frac{96p}{\varepsilon})^{6p}\sum_{m=1}^{\infty}(\mu_n/729)^{-pm} \to 0. \tag{3.12}$$

From (3.5), (3.6) and (3.12), (3.2) follows. This concludes the proof of (3.1).

By (3.1) we have

$$\left|\sum_{i=1}^{n} t_{ni}(\hat{\beta}_n)x_{ni}\right| I(\hat{\beta}_n \in D) \to 0 \quad \text{in probability, as } n \to \infty \tag{3.13}$$

and (2.7) gives, for n large, $\Delta/(\sqrt{p}\, d_n) > \mu_n^2$ (Δ is the number appearing in condition 1 of Theorem 1). So by Lemma 2 we have

$$I\left(\left|\sum_{i=1}^{n} \operatorname{sgn}(e_{ni} - x'_{ni}\hat{\beta}_n)x_{ni}\right| \geq (p+1)d_n\right) I(|\hat{\beta}_n| < \mu_n) = 0, \quad \text{a.s.}$$

Hence by (1.19)

$$\left|\sum_{i=1}^{n} \operatorname{sgn}(e_{ni} - x'_{ni}\hat{\beta}_n)x_{ni}\right| I(|\hat{\beta}_n| < \mu_n) \leq (p+1)d_n \to 0, \quad \text{a.s.} \tag{3.14}$$

From (2.12)-(2.14), (3.13) and (3.14), we have

$$\left|\sum_{i=1}^{n} \lambda_{ni}(\hat{\beta}_n)x_{ni} - \sum_{i=1}^{n} \operatorname{sgn}(e_{ni})x_{ni}\right| I(|\hat{\beta}_n| < \mu_n) \to 0, \quad \text{in probability.} \tag{3.15}$$

Since $\sup\{|x'_{ni}\hat{\beta}_n| : |\hat{\beta}_n| < \mu_n,\ 1 \leq i \leq n\} \leq \mu_n d_n \leq \mu_n^2 d_n \to 0$, we have $\lambda_{ni}(\hat{\beta}_n) = 2f(0)\big(1 + o_p(1)\big)x'_{ni}\hat{\beta}_n$ on account of condition 1 of Theorem 1, where $o_p(1) \to 0$ in probability as $n \to \infty$ uniformly for $1 \leq i \leq n$. From this, (1.18) and (3.15), we have

$$\left|2f(0)\hat{\beta}_n - \sum_{i=1}^{n} \operatorname{sgn}(e_{ni})x_{ni}\right| I(|\hat{\beta}_n| < \mu_n) \to 0, \quad \text{in probability.} \tag{3.16}$$

In view of (1.18), (1.19) and the assumptions on $\{e_{ni}\}$, it follows by Lindeberg's theorem that

$$\sum_{i=1}^{n} \operatorname{sgn}(e_{ni})\underset{\sim}{x}_{ni} \xrightarrow{\mathcal{L}} N(\underset{\sim}{0}, \underset{\sim}{I}_p). \tag{3.17}$$

From (3.16), (3.17) and Lemma 4, we obtain (1.20) (notice (2.2)). This concludes the proof of Theorem 1', hence Theorem 1.

Proof of Theorem 3. The proof differs from the above argument only in some minor details, therefore omitted.

Proof of Theorem 2. Define $\underset{\sim}{T}_n$ by (1.12), and

$$\underset{\sim}{\beta}_{n0} = \underset{\sim}{T}_n^{1/2}\beta_0, \qquad \alpha_{n0} = \sqrt{n}(\alpha_0 + \overline{\underset{\sim}{x}}_n'\underset{\sim}{\beta}_0), \qquad \underset{\sim}{\gamma}_{n0} = (\alpha_{n0}, \underset{\sim}{\beta}_{n0}')'$$

$$\underset{\sim}{x}_{ni} = \underset{\sim}{T}_n^{-1/2}(\underset{\sim}{x}_i - \overline{\underset{\sim}{x}}_n), \quad i = 1,\ldots,n,$$

$$\underset{\sim}{z}_{ni} = (1/\sqrt{n}, \underset{\sim}{x}_{ni}')', \quad i = 1,\ldots,n.$$

We transform the model (1.11) into the following form:

$$Y_i = \underset{\sim}{z}_{ni}'\underset{\sim}{\gamma}_{n0} + e_i, \quad i = 1,\ldots,n.$$

Since $\sum_{i=1}^{n}\underset{\sim}{z}_{ni}\underset{\sim}{z}_{ni}' = \underset{\sim}{I}_{p+1}$, and (1.13) guarantees that $\max_{1\le i\le n}|\underset{\sim}{z}_{ni}| \to 0$ as $n \to \infty$, Theorem 1' can be applied, and we obtain

$$2f(0)(\hat{\underset{\sim}{\gamma}}_n - \underset{\sim}{\gamma}_{n0}) \xrightarrow{\mathcal{L}} N(\underset{\sim}{0}, \underset{\sim}{I}_{p+1}), \quad \text{as} \quad n \to \infty \tag{3.18}$$

where $\hat{\underset{\sim}{\gamma}}_n = (\hat{\alpha}_{n0}, \hat{\underset{\sim}{\beta}}_{n0}')'$, and $\hat{\alpha}_{n0}$, $\hat{\underset{\sim}{\beta}}_{n0}$ are the Minimum L_1-Norm estimates of α_{n0}, $\underset{\sim}{\beta}_{n0}$, respectively. Now (3.18) implies the assertion (1.14) and also the asymptotical independence of $\hat{\alpha}_{n0}$ and $\hat{\underset{\sim}{\beta}}_{n0}$. Finally, (1.15) follows easily from what has already been proved. Theorem 2 is proved.

Remark. Wu (1987) proved that in model (1.1) the ML_1N estimate $\hat{\underset{\sim}{\beta}}_n$ is strong consistent if the following conditions are met:

1°. $\{e_i\}$ satisfies the condition stated in Theorem 3.

2°. Define $d_n = \max(1, \|x_1\|, \ldots, \|x_n\|)$ and ρ_n = the smallest eigenvalue of S_n, then $\rho_n/(d_n^2 \log n) \to \infty$, $d_n/n^c \to 0$ for some $c > 0$.

At one time it was expected that if condition 1° is replaced by $1'$: $e_1, e_2, \ldots$ are i.i.d. and e_1 has a unique median 0, the conclusion of Theorem 1 is still true. The motivation behind this conjecture is the simple case of estimating a population median by the sample median, in which the uniqueness of the population median is enough for consistency. Yet the following example shows that this is not true:

Example. In model (1.1) take $p = 1$ (β is one-dimensional), $x_n = \log n/\sqrt{n}$, $n = 1,2,3,\ldots$, $e_1, e_2, \ldots$ are i.i.d., e_1 has a density function $f(u) = |u| I(|u| < 1)$. Here $d_n = 1$, $\rho_n = S_n \sim \frac{1}{3}(\log n)^3$. Hence condition 2° is fulfilled.

In this example all conditions of Theorem 1, except that $f(0) > 0$, are met. In the course of proving Theorem 1' we have already shown this (see (3.15)). (Note that in proving (3.15) we made no use of $f(0) > 0$.)

$$\rho_n^{-1/2}\left|\sum_{i=1}^n \operatorname{sgn}(e_i)x_i - 2\sum_{i=1}^n x_i \int_0^{x_i\hat{\beta}_n} f(u)du\right| I(|\hat{\beta}_n| < 1) \to 0, \text{ in probability.} \quad (3.19)$$

Now if

$$\hat{\beta}_n \to 0, \quad \text{in probability,} \quad (3.20)$$

then since $\{x_i\}$ is bounded, from (3.19) and $f(u) = |u| I(|u| < 1)$ we have

$$\rho_n^{-1/2} \sum_{i=1}^n \operatorname{sgn}(e_i)x_i - \rho_n^{-1/2} \hat{\beta}_n^2 \sum_{i=1}^n x_i^3 \to 0, \quad \text{in probability.} \quad (3.21)$$

But by Lindeberg's theorem we have $\sqrt{\rho_n}^{-1} \sum_{i=1}^n \operatorname{sgn}(e_i)x_i \xrightarrow{\mathcal{L}} N(0,1)$, while $\sum_{i=1}^n x_i^3$ is bounded in n, $\rho_n \to \infty$ and $\hat{\beta}_n \to 0$ in probability. Thus (3.21) is impossible, which in turn implies that (3.20) is impossible.

REFERENCES

[1] AMEMIYA, T. (1982). Two stage least absolute deviations estimates. *Econemetrika* **50**. 689-711.

[2] BASSETT, G., Jr. and KOENKER, R. (1978). Asymptotic theory of least absolute error regression. *J. Amer. Statist. Assoc.* **73**, 618-622.

[3] BLOOMFIELD, P. and STEIGER, W.L. (1983). *Least Absolute Deviations.* Birkhäuser.

[4] BENNETT, G. (1962). Probability inequalities for the sums of independent random variables. *J. Amer. Statist. Assoc.* **57**, 33-45.

[5] CHARNES, A., COOPER, W.W. and FERGUSON, P.O. (1955). Optimal estimation of executive compensation by linear programming. *Mgt. Sci.* **1**, 138-151.

[6] DUPAČOVÁ, J. (1987). Asymptotic properties of restricted L_1-estimates of regression. *Statistical Data Analysis Based on the L_1-norm and Related Methods.* North-Holland.

[7] JAECKEL, L.A. (1972). Estimating regression coefficients by minimizing the dispersion of residuals. *Ann. Math. Statist.* **43**, 1449-1458.

[8] JUREČKOVÁ, J. (1971). Nonparametric estimate for regression coefficients. *Ann. Math. Statist.* **42**, 1328-1338.

[9] von NEUMANN, J. (1937). Some matrix inequalities and metrization of matrix space. *Tomsk University Review* 1, 283-300. Reprinted (1962) in *John von Neumann Collected Works* (A.H. Taub, ed.) **4**, 205-219, Pergamon, New York.

[10] WU, C.F. (1981). Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* **9**, 501-513.

[11] WU, Y. (1987). Strong consistency and exponential rate of the "Minimum L_1-Norm" estimates in linear regression models. Technical Report 87-18, Center for Multivariate Analysis, University of Pittsburgh.

APPENDIX 1

In this appendix, all notations and numbering of formula are according to Bassett and Koenker (1978), if not defined here.

Consider the model

$$y_t = x_t\beta + u_t, \quad t = 1,\ldots,T,\ldots$$

where β, one-dimensional, is the unknown parameter, u_1, u_2, ... are independent random errors with a common distribution $N(0,1)$, and

$$x_1 = 1/\sqrt{2}, \qquad x_2 = 1 + \sqrt{2}/10, \qquad x_3 = x_4 = \ldots = 1.$$

First we verify that the minimum point of the function

$$J(\beta) = \sum_{t=1}^{T} |y_t - x_t\beta|$$

is unique. For if this is not true, then owing to the convexity of $J(\beta)$, there would exist an interval $[a,b]$, $-\infty < a < b < \infty$, such that $J(c) = \inf\{J(\beta): -\infty < \beta < \infty\}$ for each $c \in (a,b)$. Choose a point $r \in (a,b)$, $r \neq y_t/x_t$, $t = 1,\ldots,T$. We should have $J'(r) = 0$, i.e.,

$$-\frac{1}{\sqrt{2}}\operatorname{sgn}(y_1 - r/\sqrt{2}) - \left(1+\frac{\sqrt{2}}{10}\right)\operatorname{sgn}\left(y_2 - \left(1+\frac{\sqrt{2}}{10}\right)r\right) - \sum_{t=3}^{T}\operatorname{sgn}(y_t - r) = 0.$$

But this is impossible, since the sum of the first two terms is an irrational number while the third term is an integer. This proves the uniqueness stated above.

Now in this model $H = \{1,2,3,\ldots,T\}$. By the choice of $\{x_t\}$, the distribution of $Z_T(\delta,1)$ is nonlattice. So according to (3.9), we should have

$$\lim_{T\to\infty} T^{1/2}\Pr\left(Z_T(\delta,1) \in C[0,1]\right) \text{ exists and not zero.} \qquad (*)$$

But

$$Z_T(\delta,1) = (\sqrt{2}+\tfrac{1}{5})\operatorname{sgn}\left(u_2 - T^{-1/2}(1+\tfrac{\sqrt{2}}{10})\delta\right) + \sum_{t=3}^{T}\sqrt{2}\operatorname{sgn}(u_t - T^{-1/2}\delta).$$

Therefore, when T is even the right-hand side, with probability one, equals to an odd multiple of $\sqrt{2}$ plus ±1/5, which is always outside [-1,1]. Consequently we have

$$P\left(Z_T(\delta,1) \in C[0,1]\right) = 0, \quad \text{for} \quad T = 2,4,6,\ldots$$

and (*) breaks down.

APPENDIX 2

In this appendix, all notations and numbering of formula are according to Amemiya (1982) if not defined here.

1. Denote by $A_1(\beta_0)$ the value of A_1 in (3.12) taken at the true parameter point β_0. We shall proceed to show that $A_1(\beta_0) \to \infty$ in probability as $T \to \infty$.

Define

$$\xi_t = \begin{cases} 1, & \text{if } |y_t - x_t'\beta_0| \le C_T^{-1} \\ 0, & \text{otherwise.} \end{cases}$$

Since $C_T = T^d$, $\frac{1}{3} < d < \frac{1}{2}$, it follows that $C_T^{-1} < T^{-1/3}$. Hence $Z_t \ge \xi_t$, and

$$\begin{aligned} A_1(\beta_0) &= 2C_T^{-1} \sum_{i=1}^{T} Z_t \log(1 + e^{-C_T|y_t - x_t'\beta_0|}) \\ &\ge 2C_T^{-1} \sum_{t=1}^{T} \xi_t \log(1 + e^{-C_T|y_t - x_t'\beta_0|}) \\ &\ge 2C_T^{-1} \log(1 + e^{-1}) \sum_{t=1}^{T} \xi_t \equiv \tilde{A}. \end{aligned}$$

Since $y_t - x_t'\beta_0 = u_t$, $t = 1,2,\ldots$ are independent and identically distributed with a common density function f which is continuous and $f(0) > 0$, it follows that there exist two positive constants h_1 and h_2 not depending on t, such that

$$h_1 C_T^{-1} \le P(\xi_t = 1) \le h_2 C_T^{-1}, \quad t = 1,2,\ldots .$$

Therefore we obtain

$$E(\tilde{A}) \ge 2\log(1 + e^{-1}) h_1 T C_T^{-2} \to \infty.$$

Here we used the fact that $C_T = T^d$ and $1/3 < d < 1/2$. Further

$$\mathrm{Var}(\tilde{A}) \leq \left(2C_T^{-1}\log(1+e^{-1})\right)^2 \sum_{t=1}^{T} E\xi_t^2$$

$$\leq \left(2C_T^{-1}\log(1+e^{-1})\right)^2 Th_2C_T^{-1}$$

$$= 4h_2\log^2(1+e^{-1})T/C_T^3 \to 0$$

by the definition of C_T given above. From $E(\tilde{A}) \to \infty$ and $\mathrm{Var}(\tilde{A}) \to 0$, we have $\tilde{A} \to \infty$ in probability as $T \to \infty$. Since $A_1(\underset{\sim}{\beta}_0) \geq \tilde{A}$, we obtain $A_1(\underset{\sim}{\beta}_0) \to \infty$ in probability as $T \to \infty$.

2. Denote by $B_1(\underset{\sim}{\beta}_0)$ the value of B_1 (in (3.22)) taken at the true parameter point $\underset{\sim}{\beta}_0$. We shall now show that $B_1(\underset{\sim}{\beta}_0) \to \infty$ in probability as $T \to \infty$.

Define ξ_t as before. Since $y_t - \underset{\sim}{x}_t'\underset{\sim}{\beta}_0 = u_t$, we have

$$B_1(\beta_0) = T^{-1/2}\sum_{t=1}^{T} I_{(|u_t|<T^{1/3}C_T^{-1})}|W_t - G_0(u_t)||x_{it}|$$

$$\geq \frac{1}{M(e+1)}T^{-1/2}\sum_{t=1}^{T}\xi_t|x_{it}|^2 \equiv \tilde{B}$$

where $M = \sup\{|\underset{\sim}{x}_t| : t = 1,2,\ldots\} < \infty$ by assumption. We have

$$E(B) \geq \frac{1}{M(e+1)}T^{-1/2}h_1C_T^{-1}\sum_{t=1}^{T}x_{it}^2$$

$$= \frac{1}{M(e+1)}h_1T^{1/2}C_T^{-1}\frac{1}{T}\sum_{t=1}^{T}x_{it}^2 \to \infty, \quad \text{as} \quad T \to \infty.$$

This is because $\sum_{t=1}^{T}x_{it}^2/T$ tends to a positive limit as $T \to \infty$ (by assumption), and that $C_T = T^d$, $1/3 < d < 1/2$. Further

$$\begin{aligned}
\mathrm{Var}(\tilde{B}) &\leq M^{-2}(e+1)^{-2}T^{-1}\sum_{t=1}^{T}|x_{it}|^4 E\xi_t^2 \\
&\leq M^{2}(e+1)^{-2}T^{-1}\sum_{t=1}^{T} E\xi_t^2 \\
&\leq M^{2}(e+1)^{-2}T^{-1}\,Th_2C_T^{-1} \to 0, \quad \text{as} \quad T \to \infty.
\end{aligned}$$

Since $E(\tilde{B}) \to \infty$ and $\mathrm{Var}(\tilde{B}) \to 0$, we have $\tilde{B} \to \infty$ in probability as $T \to \infty$. Since $B_1(\underset{\sim}{\beta}_0) \geq \tilde{B}$, the same is true for $B_1(\underset{\sim}{\beta}_0)$.